Deep learning framework for urban seismic risk assessment: A two-phase similarity algorithm for damage prediction and loss estimation across European-Mediterranean cities

Oğuz Deniz*[a], Hüseyin Pekkan[a], A. Murat Özbayoğlu[a], Onur Lenk[b], Seda Nur Kara[b], and Tacettin Yücel İpek[b]

[a]TOBB University of Economics and Technology, Söğütözü St. No:43, 06560, Ankara, Türkiye
[b]Nik System, Ortaklar Sk. Yönet Apt. No:27 D:6, Şişli 34394, İstanbul, Türkiye

ABSTRACT

We present an integrated, two-step framework for high-resolution urban seismic risk assessment and similarity-based damage prediction developed under the UDENE[1] initiative. The methodology couples an Earthquake Hazard Assessment, producing rasterized peak ground acceleration, spectral acceleration and intensity fields at 150 arcsec grid resolution using region specific Ground Motion Prediction Equation (GMPE), Boore, D. M., and Atkinson, G. M.[2] augmented by local soil information, with an Earthquake Loss Assessment that converts hazard outputs into some variety levels of building damages and economic loss estimates using a geo-cellular inventory, following Risk-UE taxonomy,[3] and the Lagomarsino–Giovinazzi[4] vulnerability relationships. To generalize predictions across contexts we develop a two-phase similarity and prediction algorithm. Phase 1 computes weighted similarity scores across Damage Loss (D3 to D5), Economic Loss, and Ground Motion using distributional and spatial measures and SSIM[5] on PGA/intensity images; weights are tunable and an optional feature importance term may be included. Phase 2 trains regression models (Random Forest,[6] Gradient Boosting[7]) on building and region-specific attributes drawn from Phase-1 clusters to predict damage classes and economic loss. Tested on 16 European–Mediterranean cities, the framework yields low mean percentage errors on held-out tests. Comprehensive sensitivity and interpretability analyses identify spatial dependency along active fault segments and reveal key vulnerability drivers. We discuss policy implications, including building-code effectiveness and an "inverse" planning scenario in which persistent high damage would motivate relocation or horizontal urban development, and demonstrate how this explainable, data-driven toolkit can support resilient urban planning and emergency preparedness.

Keywords: Urban seismic risk assessment, Earthquake loss prediction, Similarity-based region analysis, High-rise vulnerability

1. INTRODUCTION

Urban seismic risk poses a persistent and evolving challenge for contemporary cities. The combination of rapid urbanization, densification of city centers, and the growing interdependence of infrastructure systems has increased the potential for earthquakes — even of moderate magnitude — to produce outsized human, economic, and systemic impacts. Classical approaches to urban seismic risk emphasize the need to couple physical hazard characterization with an understanding of the built environment's vulnerability and the population's exposure; they also underscore that meaningful loss estimates require integration across disciplines and scales, from ground-motion physics to socio-economic exposure and post-event service functionality.[8–10] These foundational perspectives continue to guide practitioners and researchers, yet the rapid expansion of available data and computational methods creates an opportunity to rethink how hazard realism, vulnerability modeling, and prediction can be combined to support planning and policy in a more transferable and interpretable way.

Over the last decade, machine learning (ML) and deep learning (DL) techniques have emerged as powerful complements to traditional seismology and earthquake engineering. By drawing patterns from heterogeneous

*Corresponding author: Oğuz Deniz, E-mail: oguzdeniz179@gmail.com

data sources—seismic waveforms, ShakeMap-style[11] intensity rasters, building inventories, and socio-economic indicators—ML approaches have improved earthquake detection, phase picking, damage classification, fragility estimation, and rapid loss forecasting.[12,13] These methods bring clear advantages in capturing complex, non-linear relationships, enabling rapid post-event screening, and allowing for scalable ensemble approaches. At the same time, important caveats arise in their application: ML models are sensitive to data quality, can inherit biases from training sets, and often require explicit efforts to ensure interpretability and domain consistency.[14] Consequently, there is growing consensus that ML should not supplant mechanistic understanding but rather be integrated with physically grounded models to provide both predictive power and engineering interpretability.

A substantial body of recent work demonstrates how ML can augment or refine traditional vulnerability and loss-estimation frameworks. Tree-based methods, particularly Random Forest[6] and gradient-boosted ensembles,[7] have repeatedly been shown to perform well for building damage classification and regional damage mapping when combined with shaking intensity and a handful of structural features.[15–18] Machine-learning-driven fragility estimation has been shown to produce probabilistic vulnerability curves for reinforced concrete structures that better capture empirical variability than some traditional formulations.[19] At the economic scale, interpretable ensemble and AutoML frameworks have been developed to predict post-earthquake economic losses and casualty rates, using multi-source features and explanation tools such as SHAP[20] to make model outputs actionable for decision-makers.[21,22] These studies collectively illustrate that, with careful feature selection and interpretability measures, ML can substantially improve the speed and often the accuracy of damage and loss estimation. Despite this progress, an important tension remains between hazard realism and ML-driven prediction. Many operational pipelines either emphasize detailed, physics-informed ground motion modeling while relying on relatively simple vulnerability/loss translation rules, or they focus on ML-based predictive models that depend on coarse or inconsistent hazard inputs and single-city training data. This separation constrains both interpretability and the capacity to generalize across geographies. The problem is not merely methodological: for urban planners and policy-makers to use predictive tools, outputs must be traceable to physically meaningful inputs (e.g., how local soil conditions modify shaking) and standardized taxonomies (e.g., Risk-UE building classes) so that results can feed into regulatory and land-use decisions. Prior efforts have therefore advocated for hybrid frameworks that preserve engineering grounding while leveraging ML's capacity to synthesize diverse datasets.[23] Another gap lies in generalizing and transferring predictive models across cities and scenarios. Planners often need to know whether a city or district has historical or simulated analogues whose responses can guide model transfer or scenario-based policy evaluation. Although transfer learning and regional ML show promise, most are event-specific or single-city and seldom employ a tunable similarity metric that jointly captures damage, economic loss, and spatial ground-motion patterns.

Motivated by these gaps and by planning questions in dense European–Mediterranean cities, we developed an integrated two-step framework that couples high-resolution hazard modeling with interpretable ML loss prediction and a similarity-based analog search. The hazard module produces 150 arcsec raster fields of PGA, spectral acceleration, and intensity using GMPE (Boore, D. M., & Atkinson, G. M) together with local soil amplification to capture realistic spatial shaking and GMPE-related uncertainty. The loss module operates on geo-cells encoded with Risk-UE building classes and replacement costs, applying the Lagomarsino–Giovinazzi vulnerability relations to generate D3, D4 and D5 damage states/levels (D3-Substantial to Heavy Damage, D4-Very Heavy Damage, D5-Destruction) and baseline economic-loss estimates, thereby preserving engineering grounding for subsequent ML steps. The core innovation is a weighted similarity metric that fuses complementary notions of analogy: distributional comparisons of damage/loss (Wasserstein distance[24]), spatial Pearson correlations[25] of gridded counts, and image-similarity (SSIM) on PGA/intensity TIFFs,[26] with practitioner-tunable weights. After identifying analog regions or geocells, we train localized regressors on building- and region-level features to predict D3/D4/D5 counts and economic loss—combining ML's predictive power with physics-based inputs and standardized taxonomies.

Our pipeline was implemented and evaluated in 16 European–Mediterranean cities, located close to seismically active fault zones prone to earthquakes at both whole-city and geocell scales considering the correlation between seismic risk and tectonic boundaries which is evaluated as strong ($> 80\%$ of major quakes), but also secondary faults and human-induced factors contribute. We used a data-cube for fast queries, model training, and visualization. The similarity-guided, localized machine learning approach produced low mean percentage errors on held-out tests, offering actionable diagnostics for urban planning.

2. METHODOLOGY

Our methodology combines a physics-informed seismic risk assessment with a two-phase similarity algorithm for analog search and adaptive damage prediction. It proceeds in two steps: first, an Earthquake Hazard and Loss Assessment generates standardized damage and economic loss estimates; second, the similarity-based algorithm enables transferable modeling across urban contexts. The framework processes data at a high spatial resolution of 150 arcsec grid geocells (approximately 4.5×4.5 km) across 16 European–Mediterranean cities, the study considers three different earthquake scenarios for each city, with the magnitudes varying depending on the specific city's seismic risk to capture its unique vulnerability patterns. The framework's design philosophy emphasizes the integration of physics-based hazard modeling with interpretable machine learning, preserving engineering grounding while enabling cross-regional generalization. This approach addresses the critical gap between detailed hazard assessment and practical urban planning applications by maintaining traceability to physically meaningful parameters throughout the prediction pipeline.

2.1 Study Area and Dataset

The framework was developed and validated using a comprehensive dataset spanning 16 European–Mediterranean cities: Athens, Barcelona, Catania, Granada, Istanbul, Lisbon, London, Malaga, Messina, Naples, Nice, Palermo, Paris, Rome, Vienna, and Vrancea. The core dataset, generated by the Earthquake Loss Estimation Routine (ELER),[27] which was developed under the Joint Research Activity-3 (JRA3) of the NERIES project (Network of Research Infrastructures for European Seismology), funded by the European Community's 6th Framework Programme (FP6), is currently maintained by Boğaziçi University, KOERI, Earthquake Engineering Department. The routine is designed to provide rapid earthquake loss estimates for Europe using pre-defined source models and regional tectonic databases. The output data contains simulated risk analyses for the stated three earthquake magnitudes.

This dataset is structured around a geo-cellular grid resolution, with each cell containing detailed building inventories, population data, along with the simulated seismic outputs. As an example, Figure 1 illustrates the content of a GeoTIFF file, which shows the spatial distribution of damaged reinforced concrete (RC) high-rise buildings (≥ 9 stories) for D3–D5 levels induced by a deterministic earthquake scenario with $M_w = 7.4$ (epicenter: 40°53′42″N, 28°43′44″E, depth = 13 km) in the Istanbul region.

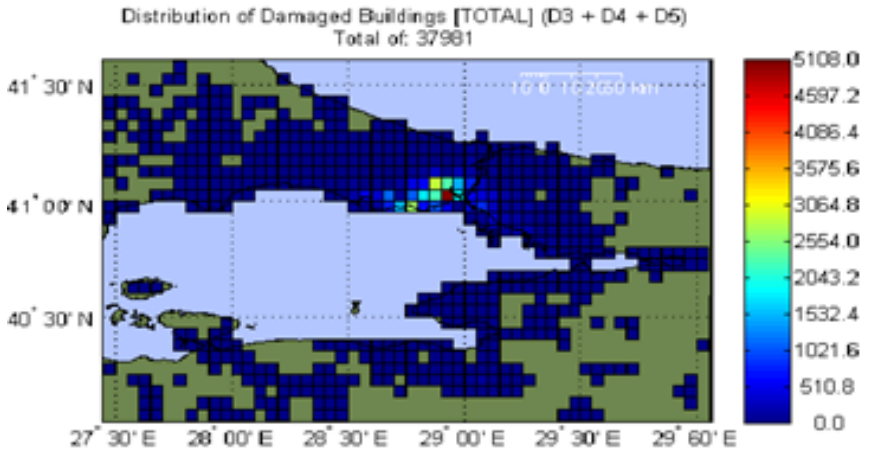


Figure 1: The spatial distribution of damaged RC high-rise buildings for D3-D5 levels (Geotiff format)

The primary data inputs for our algorithm include:

- **Damage and Economic Loss Tables:** CSV files detailing building counts in damage states D3 (Significant to Heavy Damage), D4 (Very Heavy Damage), and D5 (Destruction), alongside total economic loss per geocell.
- **Ground Motion Raster Data:** GeoTIFF (tiff) files representing the spatial distribution of ground motion intensity measures, including Peak Ground Acceleration (PGA) and Spectral Acceleration (SA) at various periods (e.g., 0.2 s and 1.0 s).
- **Regional Data:** Manual input of earthquake epicenter information (Geographic WGS84 coordinates—latitude and longitude—, seismic moment magnitude M_w, and depth of focus); ancillary text files containing parameters such as local soil conditions (V_{S30}) compiled from the USGS Global Vs30 database; active fault line data gathered from the *Global Earthquake Model Foundation (GEM)*, and the *European Seismic Hazard Model (ESHM20)*.

2.2 Earthquake Hazard and Loss Assessment

The estimation of damaged states and economic losses involves probabilistic models such as *fragility curves*, which relate ground motion intensity to the likelihood of exceeding specific damage thresholds. These curves are derived from empirical data, expert judgment, or numerical simulations, and their parameters—median ground motion values (μ) and dispersion (β)—introduce uncertainties into the analysis. For example, a lognormal probability distribution is commonly used to capture the variability in building performance, but the accuracy of such distributions relies heavily on the quality and quantity of available data. Additionally, the heterogeneity of building stock, variations in construction quality, and adherence to building codes further complicate the assessment of structural vulnerabilities. Similarly, economic loss estimation compounds the uncertainty by requiring assumptions about replacement costs, repair costs, and indirect losses such as economic disruption. These estimates are influenced by local economic conditions, infrastructure interdependencies, and the duration of service outages.

The implemented methodology, in accordance with the estimation process, mainly comprises three stages:

1. Estimation of the spatial distribution of ground motion parameters determined by means of region-specific Ground Motion Prediction Equations (GMPEs), using shear-wave velocity distributions or other regional ground values for a given earthquake magnitude, depth, and epicenter location information.
2. Estimation of building damages at different levels of complexity, commensurate with the availability of an inventory of the human-built environment.
3. Direct economic loss estimation resulting from building damages.

The initial step establishes a baseline physical risk profile for a scenario event by integrating three sequential modules: an *Earthquake Hazard Assessment*, a *Building Inventory and Vulnerability Assessment*, and an *Economic Loss Evaluation*.

2.2.1 Earthquake hazard assessment

The earthquake hazard scenarios were simulated using a deterministic approach, with epicentral information of relevant events for the stated magnitudes. The Boore and Atkinson (2008) ground motion prediction equations (GMPEs) were applied to derive intensity measures, including Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), and Spectral Acceleration (SA) at 0.2 s, 1.0 s, and 30 s. Local site conditions were incorporated through a surface correction function dependent on V_{S30} values, thereby adjusting the hazard field for near-surface amplification. To ensure consistency with vulnerability models, ground motions were converted into macroseismic intensity using the *European Macroseismic Scale (EMS-98; Grünthal, 1998)*.[28] This facilitated application of the observed-damage-based vulnerability method, often referred to as the macroseismic method[4,29] (Giovinazzi & Lagomarsino, 2004, 2005).

This approach expresses the expected mean damage grade (μ_D) as a function of seismic intensity (I), the building's vulnerability index (V), and ductility index (Q). The basic expression adopted is given in Equation 1.

$$\mu_D = 2.5 \left[1 + \tanh \left(\frac{I + 6.5V - 13.1}{2.3} \right) \right] , \tag{1}$$

where μ_D is the mean damage grade (ranging from 0 = no damage to 5 = destruction), I is the macroseismic intensity (EMS-98), V is the building vulnerability index (assigned by structural type, typically between 0.5–1.0), and Q is the ductility index (modulating the slope of the damage–intensity function).

From this function, the probability of a building reaching each EMS-98 damage grade D_k ($k = 0, \ldots, 5$) is obtained using a **binomial probability distribution**:

$$P(D_k) = \binom{5}{k} \mu_D^k (1 - \mu_D)^{5-k} . \tag{2}$$

The expected number of buildings in each damage state is then calculated as:

$$N_{\mathrm{DS}_i} = P(\mathrm{DS}_i) \cdot N_{\mathrm{total}} , \tag{3}$$

where N_{DS_i} is the expected number of buildings in damage state DS_i, $P(\mathrm{DS}_i)$ is the probability of being in damage state DS_i, and N_{total} is the total number of buildings in the inventory.

These relationships enable the translation of intensity fields into probabilistic distributions of structural damage across the exposed building stock. The resulting data is displayed as rasterized fields (PGA, $\mathrm{PSA}_{0.2}$, PSA_{30}, and intensity) and stored in GeoTIFF files, enabling subsequent spatial analysis and visualization of seismic hazard patterns.

2.2.2 Building Inventory and Vulnerability Assessment

The building inventory for the study area was constructed using proxy data sources. In the absence of a detailed regional building database, building exposure was approximated through the *CORINE Land Cover dataset* (European Environment Agency, 1999) combined with population distribution data from *LandScan*. The method assumes a correlation between population density, land cover class, and the distribution of residential and non-residential buildings. This proxy inventory was then spatially aggregated into grid cells of 150 arcseconds (approximately 4.5 km resolution), forming the default exposure dataset for the analysis.

For the classification of buildings, the *RISK-UE taxonomy* (2001–2004) was adopted. This taxonomy provides a harmonized European framework that categorizes buildings according to their structural typologies, materials, and construction practices. It enables consistent linkage between building classes and vulnerability functions developed for seismic risk assessment. By applying this taxonomy, the regional stock was represented in terms of dominant structural types, allowing for systematic vulnerability assessment across the study area.

Vulnerability relationships were derived from the models proposed by Lagomarsino and Giovinazzi (2006), which link EMS-98 intensity levels to expected damage grades for different building typologies. These relationships are probabilistic functions that define the likelihood of buildings reaching various states of damage (slight, moderate, extensive, and complete) under a given ground motion intensity. The application of these vulnerability curves ensured compatibility with the hazard inputs expressed in EMS-98 intensities, thereby enabling coherent building damage estimation. The methodology evaluates three critical damage states: D3 (Significant to Heavy Damage), D4 (Very Heavy Damage), and D5 (Destruction), which represent progressively severe levels of structural compromise.

The vulnerability calculation process operates at the geocell level, with each cell containing detailed building inventory data including construction type distribution, total building counts, and population density. The damage assessment considers multiple building height categories (RC L: low-rise, RC M: mid-rise, RC H: high-rise reinforced concrete structures) to capture the differential seismic response of buildings with varying dynamic characteristics.

2.2.3 Economic loss evaluation

Economic losses were computed directly from the building damage distributions estimated through the macroseismic vulnerability method. In ELER's Level 1 framework, losses are expressed as the ***Expected Loss Ratio (ELR)***, which combines the probabilities of damage states with their associated repair cost factors. For each building class c in grid cell j, the ELR is defined as:

$$\mathrm{ELR}_{c,j} = \sum_{k=0}^{5} P(D_{k,c,j}) \cdot R_k \,, \tag{4}$$

where $P(D_{k,c,j})$ is the probability that buildings of class c in cell j are in damage state k, and R_k is the repair cost ratio associated with damage state k (ranging from 0 for no damage to ~ 1.0 for complete destruction).

The direct economic loss for each grid cell is then calculated as:

$$\mathrm{Loss}_j = \sum_{c} N_{c,j} \cdot V_c \cdot \mathrm{ELR}_{c,j} \,, \tag{5}$$

where $N_{c,j}$ is the number of buildings of class c in grid cell j, V_c is the average replacement cost per building of class c, and $\mathrm{ELR}_{c,j}$ is the expected loss ratio for class c in cell j.

This formulation, adapted from Giovinazzi & Lagomarsino (2006) and implemented in ELER (*RISK-UE*, 2004), provides a quantitative framework for converting structural damage into direct monetary losses. The resulting loss distribution, aggregated at the 150 arcsec grid scale, reflects the spatial pattern of potential economic impacts. While this method relies on proxy exposure data and approximate repair cost ratios, it provides a consistent and scalable approach to evaluate direct losses at the European scale. Future refinements may incorporate country-specific replacement values or calibration with insurance and post-disaster datasets.

2.3 Two-Phase Similarity and Prediction Algorithm

The core innovation of our work is a two-phase algorithm designed to first identify regions with similar seismic risk profiles and then train customized prediction models for them. This structure stabilizes the influence of seismic and geographical variables in the first phase, allowing the second phase to focus on predicting vulnerability based on building-specific features.

2.3.1 Data preprocessing and structure

The algorithm operates on preprocessed datasets extracted from the comprehensive earthquake simulation results. The preprocessing stage selects relevant data components from the raw NiK System outputs: damage classification tables (D3, D4, D5), total economic loss tables, ground motion raster files, and site condition data. This preprocessing ensures data consistency and compatibility across the 16-city dataset while maintaining the high-resolution geocell structure essential for spatial analysis.

2.3.2 Phase 1: weighted similarity assessment

The objective of Phase 1 is to identify the top-k regions or geocells most similar to a given query input. This is accomplished by calculating a custom, multi-component **Overall Similarity Score**. This score is a weighted aggregation of three primary similarity metrics: Damage Loss Similarity, Economic Loss Similarity, and Ground Motion Similarity.

Damage Loss Similarity
For each damage class (D3, D4, D5), the algorithm computes a composite Damage Similarity score through four complementary metrics:

1) Distribution Similarity: Quantifies the statistical similarity of damage count distributions using the Wasserstein distance metric applied to the expected damage ratios across geocells. The expected damage ratio in each cell represents the number of buildings projected to be temporarily unusable after an earthquake.

This is typically derived from the estimated count of buildings with moderate or higher damage levels, where even partial structural damage can render a building unsuitable for immediate occupancy.

This measure captures both the magnitude and the distributional shape of damage patterns between regions. The Wasserstein distance is defined as:

$$\mathcal{W}_1(P,Q) = \inf_{\gamma\in\Gamma(P,Q)} \int_{\mathbb{R}\times\mathbb{R}} |x-y|\,\mathrm{d}\gamma(x,y), \tag{6}$$

where $\Gamma(P,Q)$ denotes the set of all joint distributions with marginals P and Q. For empirical (discrete) distributions this corresponds to the usual Earth Mover's Distance computed on the damage-count support.

Denoting $W = \mathcal{W}_1(P,Q)$, we convert the numeric distance to a similarity score on $[0,1]$ via a monotonic mapping. Two practical choices are:

$$S_{\text{distribution}} = 1 - \frac{W - W_{\min}}{W_{\max} - W_{\min}}, \qquad W_{\min} \le W \le W_{\max}, \tag{7}$$

for min–max normalization, and

$$S_{\text{distribution}} = \exp(-\alpha W), \qquad \alpha > 0, \tag{8}$$

for the exponential kernel, where $W_{\min}$ and $W_{\max}$ are the minimum and maximum Wasserstein distances observed (or estimated) across the search space and α is a tunable scale parameter. The mapping strategy is configurable in the pipeline; the default is min–max normalization, yielding a bounded Distribution Similarity $S_{\text{distribution}} \in [0,1]$.

2) Spatial Similarity: Evaluates the geographic correlation of damage patterns using Pearson correlation coefficients calculated on gridded damage counts with spatial coordinates. The spatial analysis employs a 50-cell grid resolution to balance computational efficiency with spatial detail preservation.

The vectors $\mathbf{c}^{(q)}, \mathbf{c}^{(\text{cand})} \in \mathbb{R}^M$, where M is the number of grid cells after aggregation.

The spatial similarity is computed from Pearson's correlation as

$$S_{\text{spatial}} = \rho\big(\mathbf{c}^{(q)}, \mathbf{c}^{(\text{cand})}\big), \tag{9}$$

with

$$\rho\big(\mathbf{c}^{(q)}, \mathbf{c}^{(\text{cand})}\big) = \frac{\sum_{i=1}^{M} \big(c_i^{(q)} - \bar{c}^{(q)}\big)\big(c_i^{(\text{cand})} - \bar{c}^{(\text{cand})}\big)}{\sqrt{\sum_{i=1}^{M} \big(c_i^{(q)} - \bar{c}^{(q)}\big)^2}\,\sqrt{\sum_{i=1}^{M} \big(c_i^{(\text{cand})} - \bar{c}^{(\text{cand})}\big)^2}}, \tag{10}$$

where $\bar{c}^{(q)}$ and $\bar{c}^{(\text{cand})}$ are the sample means of the respective vectors.

When a bounded similarity in $[0,1]$ is required, $\rho \in [-1,1]$ can be rescaled as

$$S_{\text{spatial}}^{[0,1]} = \frac{\rho\big(\mathbf{c}^{(q)}, \mathbf{c}^{(\text{cand})}\big) + 1}{2}. \tag{11}$$

The aggregation window G (default 50) and the choice whether to use raw counts or loss ratios are tunable parameters in the pipeline.

3) Population Similarity: Assesses demographic exposure through population-to-building ratios, capturing the relationship between human exposure and built environment characteristics. The population-to-building ratio in geocell c is defined as:

$$r_c = \frac{P_c}{B_c + \varepsilon}, \qquad \varepsilon > 0, \tag{12}$$

where P_c denotes the population in geocell c and B_c denotes the number of buildings. We set $\varepsilon = 1$ by default to avoid division by zero when $B_c = 0$.

For a query region and a candidate region, let $r^{(q)}$ and $r^{(\text{cand})}$ denote their respective ratios. The scalar population similarity is computed as the inverse normalized absolute difference:

$$S_{\text{population}} = 1 - \frac{|\, r^{(q)} - r^{(\text{cand})}\,|}{R_{\max} - R_{\min}}, \qquad R_{\min} \le |r^{(q)} - r^{(\text{cand})}| \le R_{\max}, \tag{13}$$

with values clipped to $S_{\text{population}} \in [0,1]$. Here $R_{\min}$ and $R_{\max}$ denote the minimum and maximum absolute differences observed (or estimated) across the candidate set.

When ratios are available as aligned vectors over M grid cells, $\mathbf{r}^{(q)}, \mathbf{r}^{(\text{cand})} \in \mathbb{R}^M$, the similarity is defined using the mean absolute difference:

$$S_{\text{population}} = 1 - \frac{\frac{1}{M}\sum_{i=1}^{M} |\, r_i^{(q)} - r_i^{(\text{cand})}\,|}{R_{\max} - R_{\min}}, \tag{14}$$

again clamped to $[0,1]$.

As an alternative, one may use an exponential kernel,

$$S_{\text{population}} = \exp\big(-\beta |r^{(q)} - r^{(\text{cand})}|\big), \qquad \beta > 0, \tag{15}$$

where β is a tunable scale parameter. The mapping strategy and the choice of scalar versus vector aggregation are configurable in the pipeline, with min–max normalization as the default.

4) Building Similarity: Building similarity compares the total building inventory characteristics between regions, accounting for the influence of building stock density on damage patterns. For scalar inventories, it is defined as:

$$S_{\text{building}} = 1 - \frac{|B^{(q)} - B^{(\text{cand})}|}{\max\big(B^{(q)}, B^{(\text{cand})}, \varepsilon\big)}, \qquad \varepsilon = 1 \text{ (default)}, \tag{16}$$

which ensures $S_{\text{building}} \in [0,1]$.

For gridded inventories $\mathbf{b}^{(q)}, \mathbf{b}^{(\text{cand})} \in \mathbb{R}^M$, the similarity extends to

$$S_{\text{building}} = 1 - \frac{1}{M}\sum_{i=1}^{M} \frac{|b_i^{(q)} - b_i^{(\text{cand})}|}{\max\big(b_i^{(q)}, b_i^{(\text{cand})}, \varepsilon\big)}. \tag{17}$$

As an alternative, one may use Pearson correlation rescaled to $[0,1]$,

$$S_{\text{building}} = \frac{\rho + 1}{2}, \tag{18}$$

where ρ is the Pearson correlation coefficient between $\mathbf{b}^{(q)}$ and $\mathbf{b}^{(\text{cand})}$.

Composite Damage Similarity:

The composite damage similarity for each damage class is defined as:

$$S_{D_i}(d) = w_1\, S_{\text{distribution}}(d) + w_2\, S_{\text{spatial}}(d) + w_3\, S_{\text{population}}(d) + w_4\, S_{\text{building}}(d), \qquad d \in \{D3, D4, D5\}, \tag{19}$$

where $S_{\text{distribution}}$, S_{spatial}, $S_{\text{population}}$, and S_{building} are the component similarities.

With the default weights

$$w_1 = 0.40, \quad w_2 = 0.30, \quad w_3 = 0.15, \quad w_4 = 0.15.$$

Equation (19) is evaluated separately for each damage class $D3$, $D4$, and $D5$, producing three composite similarity values.

Economic Loss Similarity

Economic loss similarity consists of two complementary measures: a base similarity that compares the overall magnitude of losses between regions, and a spatial loss-ratio similarity that evaluates the distribution of losses normalized by exposure.

Base similarity, S_{base}, applies the same distributional approach used for damage assessment to economic loss values, comparing the total economic loss values of a query region Q and a candidate region C.

$$S_{\text{base}}(Q, C) = 1 - \frac{|L_Q - L_C|}{\max(L_Q, L_C)}\,, \tag{20}$$

where L_Q and L_C denote the total economic losses of the query and candidate regions, respectively. This formulation ensures a bounded similarity score between 0 and 1.

Spatial loss ratio similarity, $S_{\text{spatialLossRatio}}$, incorporates geographic weighting by analyzing the ratio of economic losses between compared regions to capture how well the spatial distribution of normalized losses aligns between the two regions.

$$S_{\text{spatialLossRatio}}(Q, C) = \text{corr}\left(\frac{L_{Q,i}}{E_{Q,i}}, \frac{L_{C,i}}{E_{C,i}}\right)\,, \tag{21}$$

where $L_{Q,i}$ and $L_{C,i}$ represent the economic losses in geocell i of regions Q and C, and $E_{Q,i}$, $E_{C,i}$ are the corresponding exposure metrics (e.g., total building replacement cost). The correlation is typically measured using Pearson's r.

The composite Economic Loss Similarity is calculated as:

$$S_{\text{Econ}} = w_1 \cdot S_{\text{base}} + w_2 \cdot S_{\text{spatialLossRatio}}\,, \tag{22}$$

where the default weights are $w_1 = 0.7$ and $w_2 = 0.3$.

Ground Motion Similarity

Ground motion comparison employs the *Structural Similarity Index Measure* (SSIM) applied to the GeoTIFF raster files containing **PGA**, **PSA02**, and **PSA30** values. SSIM provides a perceptually-motivated similarity measure that captures both intensity and spatial pattern similarities in ground motion fields.

The ground motion similarity is defined as

$$S_{\text{GM}} = w_1 \cdot \text{SSIM}_{\text{PGA}} + w_2 \cdot \text{SSIM}_{\text{PSA02}} + w_3 \cdot \text{SSIM}_{\text{PSA30}}\,, \tag{23}$$

where the default weights are $w_1 = 0.33$, $w_2 = 0.33$, and $w_3 = 0.34$.

Overall Similarity Score

The final similarity assessment integrates all component similarities through a weighted sum with practitioner-tunable weights:

$$S_{\text{Overall}} = W_{D3}S_{D3} + W_{D4}S_{D4} + W_{D5}S_{D5} + W_{\text{Econ}}S_{\text{Econ}} + W_{\text{GM}}S_{\text{GM}}\,, \tag{24}$$

where the default weights are

$$W_{D3} = 0.15, \quad W_{D4} = 0.20, \quad W_{D5} = 0.25, \quad W_{\text{Econ}} = 0.30, \quad W_{\text{GM}} = 0.10\,.$$

These weights prioritize severe damage states and economic impacts while maintaining sensitivity to ground motion patterns. An optional feature importance term, derived from regression tree analysis, can also be incorporated to enhance similarity assessment accuracy.

Operational Modes

Phase 1 operates in two distinct modes to support different analytical requirements:

City-wide Mode compares entire urban regions as integrated units, evaluating all geocells within each city collectively. This mode identifies cities with similar overall seismic response characteristics and is particularly useful for regional planning and policy transfer applications.

Geocell-based Mode performs individual geocell comparisons, enabling fine-scale similarity assessment and localized prediction. This mode supports neighborhood-level analysis and detailed urban planning applications. For both modes, the algorithm returns the k most similar regions (where k is user-specified) ranked by Overall Similarity Score, providing a flexible framework for similarity-based analysis at multiple scales.

2.3.3 Phase 2: localized damage and loss prediction

Phase 2 leverages the similarity clusters identified in Phase 1 to train specialized regression models for damage and economic loss prediction. This approach capitalizes on the principle that regions with similar seismic response characteristics will benefit from localized predictive models trained on relevant analog data.

Feature Selection Strategy

The model training protocol addresses the challenge of heterogeneous feature availability across cities through a dynamic feature selection approach. Since Phase 1 calculations predominantly utilize earthquake parameters and geographic features, the influence of these variables is captured within the similarity-based clustering process. Phase 2 models therefore focus on building-specific and regional characteristics that vary within similarity clusters. The feature selection process filters out variables already incorporated in Phase 1 similarity calculations: geocell identifiers, geometric data, regional identifiers, demographic variables, building inventory totals, and ground motion intensity. The remaining features consist primarily of building type distributions and local characteristics that influence vulnerability within seismically similar regions.

Training Data Composition

Training datasets are constructed from the geocells identified as similar in Phase 1, ensuring that models are trained on regions with comparable seismic response characteristics. This similarity-guided training approach improves model relevance and prediction accuracy compared to global models trained on heterogeneous regional data. The training process adapts to the specific feature availability within each similarity cluster, automatically handling missing building type data and ensuring robust model performance across diverse urban contexts. This adaptive approach enables the framework to operate effectively across cities with varying building inventory detail and classification systems.

Model Training and Implementation

The prediction system employs ensemble methods specifically selected for their interpretability and performance with heterogeneous feature sets. Two model types are implemented:

1) Random Forest Regressor: Configured with hyperparameters $n\_estimators = 100$, $max\_depth = 20$ and $min\_samples\_split = 5$. This configuration balances model complexity with interpretability while providing robust performance across diverse feature distributions.

2) Gradient Boosting Regressor: Configured with hyperparameters $n\_estimators = 100$, $learning\_rate = 0.1$ and $max\_depth = 5$. The gradient boosting approach provides complementary modeling capabilities, particularly for capturing non-linear relationships between building characteristics and damage outcomes.

Separate models are trained for each damage class (D3, D4, D5) and economic loss prediction, resulting in four predictive models per similarity cluster. The models are persistently stored to enable efficient inference and reuse across queries.

2.3.4 Operational Framework

To operationalize this methodology, we designed a framework centered on a data cube. This serves as a centralized repository for the entire multi-city dataset ('world data') and allows for efficient slicing and querying. When a user defines a query region or geocell, the system executes the two-phase algorithm within this environment. The outputs—including the list of similar regions, similarity scores, and damage/loss predictions—are then passed to an integrated front-end interface for visualization and analysis, enabling interactive exploration by urban planners, engineers, and policymakers (see Fig. 2).

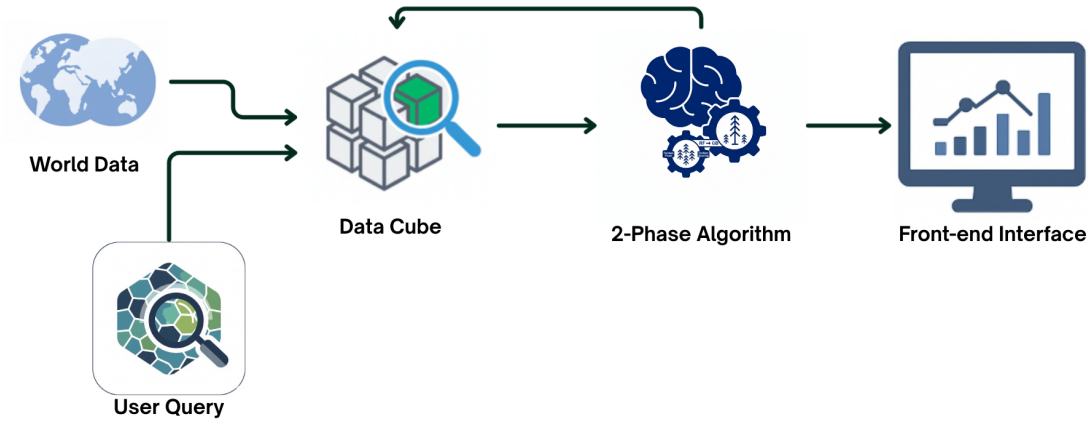


Figure 2: Workflow of the operational framework, illustrating the data cube, user query, two-phase algorithm, and front-end output.

3. RESULTS

The two-phase similarity and prediction framework was evaluated across 16 European-Mediterranean cities, with each city analyzed under three earthquake magnitude scenarios specifically designed for their unique seismic risk profiles.

3.1 Phase 1 Results: Similarity Assessment and Regional Clustering

The primary objective of Phase 1 was to quantify the similarity between different urban seismic scenarios to identify meaningful analogs for transferable risk prediction. The algorithm was applied across all 16 cities for three distinct magnitude scenarios. The top-five most similar scenarios for each query are presented in Tables 1 (lowest magnitude), 2 (middle magnitude), and 3 (highest magnitude).

City codes & Table Format: ATH = Athens, BCN = Barcelona, CTA = Catania, GRX = Granada, IST = Istanbul, LIS = Lisbon, LON = London, AGP = Malaga, MES = Messina, NAP = Naples, NCE = Nice, PMO = Palermo, PAR = Paris, ROM = Rome, VIE = Vienna, VRC = Vrancea. Each cell lists the top-5 analogs in the format "CodeMagn. (similarity)".

3.1.1 Similarity patterns across magnitude scenarios

A dominant pattern across all magnitudes is the high degree of intra-city similarity. In most cases, the scenario most similar to a given city is another scenario in the same city at a different magnitude. Table 1 shows the similarity results for the lowest-magnitude scenario per city, where self-similarity is strongest. Eleven of sixteen cities have higher-magnitude scenarios among the top-2 analogs, highlighting internal consistency in vulnerability patterns. Notable examples include Barcelona (BCN6.5 as top analog for BCN6.0, 0.7741), Granada (GRX6.5 for GRX6.0, 0.7026), and Vrancea (VRC6.5 for VRC6.0, 0.7223).

Table 1: Magnitude-1 (lowest magnitude per city) similarity results.

City	Scenario	Top-1	Top-2	Top-3	Top-4	Top-5
ATH	6.0	NAP6.0 (0.6733)	ATH6.5 (0.6646)	VIE6.1 (0.6416)	ATH7.0 (0.6380)	MES6.0 (0.6267)
BCN	6.0	BCN6.5 (0.7741)	AGP6.5 (0.7382)	BCN6.8 (0.7188)	AGP7.0 (0.6968)	MES6.0 (0.6468)
CTA	6.0	MES6.0 (0.6836)	VRC6.5 (0.6820)	GRX6.5 (0.6723)	AGP6.5 (0.6647)	BCN6.0 (0.6434)
GRX	6.0	GRX6.5 (0.7026)	AGP6.0 (0.6167)	GRX7.4 (0.6050)	VRC6.0 (0.5857)	VRC6.5 (0.5451)
IST	6.0	IST6.5 (0.6355)	LIS6.0 (0.5982)	PMO6.0 (0.5902)	PAR4.9 (0.5591)	MES6.0 (0.5563)
LIS	6.0	LIS6.5 (0.6342)	IST6.0 (0.5982)	MES6.0 (0.5702)	PAR4.9 (0.5609)	NAP6.0 (0.5593)
LON	6.0	LON5.8 (0.7316)	LON6.5 (0.6414)	PAR6.0 (0.6335)	ATH6.5 (0.5967)	PAR6.5 (0.5749)
AGP	6.0	GRX6.0 (0.6167)	GRX6.5 (0.6027)	AGP6.5 (0.5992)	VRC6.5 (0.5538)	VRC6.0 (0.5473)
MES	6.0	MES6.5 (0.7049)	CTA6.0 (0.6836)	PMO6.0 (0.6514)	AGP6.5 (0.6501)	BCN6.0 (0.6468)
NAP	6.0	NAP6.5 (0.6901)	LIS6.5 (0.6883)	ATH6.0 (0.6733)	IST6.5 (0.6529)	PMO6.0 (0.6243)
NCE	6.0	NCE6.5 (0.6878)	NCE6.8 (0.6575)	BCN6.5 (0.6396)	PMO6.5 (0.6316)	BCN6.8 (0.6283)
PMO	6.0	PMO6.5 (0.7115)	IST6.5 (0.6622)	PMO7.0 (0.6543)	MES6.0 (0.6514)	BCN6.0 (0.6444)
PAR	4.9	VRC6.5 (0.5893)	MES6.0 (0.5871)	CTA6.0 (0.5626)	LIS6.0 (0.5609)	VRC6.0 (0.5605)
ROM	6.0	VIE6.1 (0.6271)	IST6.5 (0.6232)	BCN6.5 (0.6230)	ATH6.0 (0.6226)	MES6.5 (0.6210)
VIE	6.1	VIE6.5 (0.6825)	ATH6.0 (0.6416)	MES6.5 (0.6284)	ROM6.0 (0.6271)	VIE7.0 (0.6251)
VRC	6.0	VRC6.5 (0.7223)	GRX6.0 (0.5857)	VRC7.4 (0.5803)	PAR4.9 (0.5605)	AGP6.0 (0.5473)

Table 2: Magnitude-2 (middle magnitude per city) similarity results.

City	Scenario	Top-1	Top-2	Top-3	Top-4	Top-5
ATH	6.5	ATH6.0 (0.6646)	ATH7.0 (0.6642)	NAP6.5 (0.6117)	LON6.0 (0.5967)	VIE6.1 (0.5796)
BCN	6.5	BCN6.0 (0.7741)	BCN6.8 (0.7673)	AGP6.5 (0.6748)	AGP7.0 (0.6612)	NCE6.0 (0.6396)
CTA	6.5	CTA6.0 (0.4910)	CTA7.4 (0.4771)	BCN6.0 (0.4356)	GRX7.4 (0.4346)	PMO6.0 (0.4338)
GRX	6.5	GRX7.4 (0.7314)	GRX6.0 (0.7026)	AGP6.5 (0.6951)	CTA6.0 (0.6723)	VRC6.5 (0.6636)
IST	6.5	PMO6.0 (0.6622)	NAP6.0 (0.6529)	PMO6.5 (0.6478)	IST7.4 (0.6470)	IST6.0 (0.6355)
LIS	6.5	NAP6.0 (0.6883)	LIS7.4 (0.6682)	LIS6.0 (0.6342)	IST6.5 (0.6312)	ATH6.0 (0.6172)
LON	5.8	LON6.0 (0.7316)	PAR6.0 (0.6691)	LON6.5 (0.6356)	NCE6.0 (0.6132)	VIE6.1 (0.5925)
AGP	6.5	AGP7.0 (0.7463)	BCN6.0 (0.7382)	GRX6.5 (0.6951)	BCN6.5 (0.6748)	CTA6.0 (0.6647)
MES	6.5	MES6.0 (0.7049)	MES7.4 (0.6588)	VIE6.1 (0.6284)	ROM6.0 (0.6210)	ATH6.0 (0.6202)
NAP	6.5	NAP6.0 (0.6901)	NAP7.2 (0.6535)	LIS7.4 (0.6310)	ATH6.5 (0.6117)	LIS6.5 (0.5994)
NCE	6.5	NCE6.8 (0.7047)	NCE6.0 (0.6878)	PMO7.0 (0.5873)	IST7.4 (0.5830)	BCN6.8 (0.5744)
PMO	6.5	PMO6.0 (0.7115)	PMO7.0 (0.7037)	IST6.5 (0.6478)	NCE6.0 (0.6316)	PAR6.0 (0.5951)
PAR	6.0	LON5.8 (0.6691)	PAR6.5 (0.6606)	LON6.0 (0.6335)	VRC7.4 (0.6034)	NCE6.0 (0.6024)
ROM	6.5	ROM6.7 (0.7341)	ROM6.0 (0.6085)	VIE6.5 (0.5748)	BCN6.8 (0.5725)	CTA7.4 (0.5589)
VIE	6.5	VIE6.1 (0.6825)	VIE7.0 (0.6641)	NAP6.5 (0.5872)	ATH6.5 (0.5795)	ROM6.5 (0.5748)
VRC	6.5	VRC6.0 (0.7223)	VRC7.4 (0.7179)	CTA6.0 (0.6820)	GRX6.5 (0.6636)	MES6.0 (0.6287)

Table 2 demonstrates the evolution of similarity patterns at intermediate magnitude levels, where cross-regional relationships become more prominent while maintaining some self-similarity characteristics. Cities like Barcelona and Granada continue to show strong internal coherence (BCN6.0 and BCN6.8 ranking highly for BCN6.5), while others develop stronger external analog networks. Istanbul's pattern exemplifies this transition, with Palermo (PMO6.0) emerging as the top analog (0.6622 similarity) rather than its own magnitude variants.

Table 3 shows the similarity landscape for the highest-magnitude scenarios, where universal vulnerability patterns outweigh local variations. The persistence of strong self-similarity in some cities (AGP7.0 as top analog for AGP6.5 with 0.7463, Rome ROM6.5 for ROM6.7 with 0.7341) contrasts with the emergence of diverse analog networks in others, indicating different vulnerability progression patterns across the study regions.

Table 3: Magnitude-3 (highest magnitude per city) similarity results.

City	Scenario	Top-1	Top-2	Top-3	Top-4	Top-5
ATH	7.0	ATH6.5 (0.6642)	ATH6.0 (0.6380)	VIE7.0 (0.5748)	VIE6.5 (0.5541)	MES7.4 (0.5507)
BCN	6.8	BCN6.5 (0.7673)	BCN6.0 (0.7188)	AGP7.0 (0.6342)	NCE6.0 (0.6283)	AGP6.5 (0.6267)
CTA	7.4	CTA6.0 (0.6065)	PMO7.0 (0.5714)	ROM6.7 (0.5591)	ROM6.5 (0.5589)	ATH6.5 (0.5558)
GRX	7.4	GRX6.5 (0.7314)	AGP7.0 (0.6861)	BCN6.0 (0.6190)	AGP6.5 (0.6178)	NAP6.0 (0.6062)
IST	7.4	IST6.5 (0.6470)	PMO7.0 (0.5845)	NCE6.5 (0.5830)	NCE6.8 (0.5623)	PMO6.5 (0.5614)
LIS	7.4	LIS6.5 (0.6682)	NAP6.5 (0.6310)	NAP6.0 (0.5793)	NAP7.2 (0.5666)	ATH6.5 (0.5608)
LON	6.5	LON6.0 (0.6414)	LON5.8 (0.6356)	PAR6.5 (0.5879)	NCE6.8 (0.5477)	PAR6.0 (0.5394)
AGP	7.0	AGP6.5 (0.7463)	BCN6.0 (0.6968)	GRX7.4 (0.6861)	BCN6.5 (0.6612)	BCN6.8 (0.6342)
MES	7.4	MES6.5 (0.6588)	MES6.0 (0.6427)	CTA7.4 (0.5531)	ROM6.7 (0.5527)	ATH7.0 (0.5507)
NAP	7.2	NAP6.5 (0.6535)	NAP6.0 (0.6096)	LIS7.4 (0.5666)	ATH7.0 (0.5332)	VIE7.0 (0.5217)
NCE	6.8	NCE6.5 (0.7047)	NCE6.0 (0.6575)	IST7.4 (0.5623)	PMO7.0 (0.5605)	LON6.5 (0.5477)
PMO	7.0	PMO6.5 (0.7037)	PMO6.0 (0.6543)	NCE6.0 (0.6039)	NCE6.5 (0.5873)	IST7.4 (0.5845)
PAR	6.5	PAR6.0 (0.6606)	LON6.5 (0.5879)	LON6.0 (0.5749)	LON5.8 (0.5632)	NCE6.5 (0.5509)
ROM	6.7	ROM6.5 (0.7341)	ROM6.0 (0.5917)	CTA7.4 (0.5591)	BCN6.8 (0.5532)	MES7.4 (0.5527)
VIE	7.0	VIE6.5 (0.6641)	VIE6.1 (0.6251)	ATH7.0 (0.5748)	ATH6.5 (0.5325)	PAR6.5 (0.5224)
VRC	7.4	VRC6.5 (0.7179)	NAP6.0 (0.6108)	ATH6.0 (0.6072)	ROM6.0 (0.6060)	VIE6.1 (0.6036)

3.1.2 Regional clustering and cross-national relationships

Beyond self-similarity, the results reveal geographically and tectonically coherent clusters. Cities in similar seismic and urban contexts often appear as top analogs, consistent with engineering and seismological expectations. On the Iberian Peninsula, Barcelona, Granada, and Málaga form a coherent vulnerability group across all magnitudes, exemplified by the reciprocal high rankings between Granada and Málaga (GRX6.0 ranking AGP6.0 at 0.6167, AGP6.0 ranking GRX6.0 at 0.6167; Table 1). Southern Italian cities—Catania, Messina, Naples, and Palermo—also show strong mutual affinity, with Catania–Messina particularly stable (MES6.0 consistently the top analog for CTA6.0 at 0.6836). These regional clusters likely reflect shared construction practices, vulnerability profiles, and crustal properties, highlighting the potential for knowledge transfer and generalized mitigation strategies (see Fig. 3).

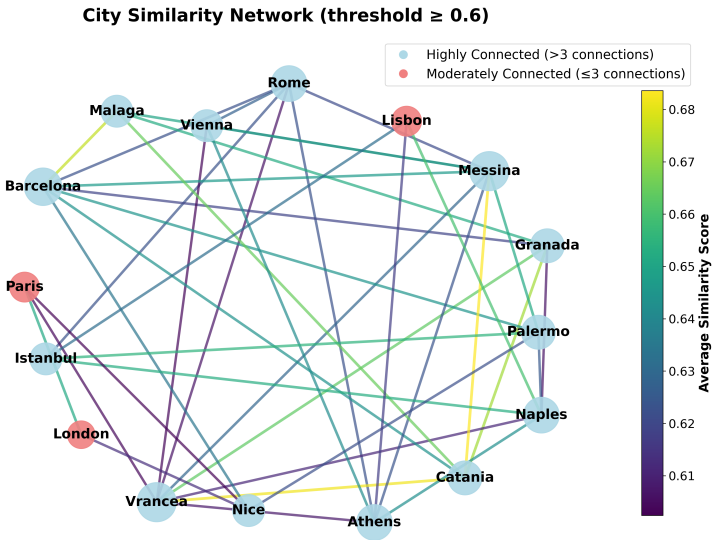


Figure 3: Similarity network of cities (threshold $\geq$ 0.6), where node colors denote connectivity and edge thickness indicates similarity strength.

At the same time, outlier patterns emerge: Rome frequently clusters with Central European cities like Vienna rather than Italian counterparts, highlighting the influence of urban development patterns and geological conditions beyond national boundaries. Vienna itself represents an especially versatile profile, ranking highly as an analog for multiple diverse cities—including Athens (VIE6.1 ranking second for ATH6.0 at 0.6416) and Rome (VIE6.1 as top analog for ROM6.0 at 0.6271)—demonstrating that its seismic vulnerability characteristics extend beyond the Central European context.

3.1.3 Magnitude-dependent similarity, vulnerability transitions, and transferability

The Phase 1 similarity assessment reveals clear, magnitude-dependent structure in urban vulnerability patterns and strong potential for transferability into Phase 2. Some cities (e.g., Barcelona, Granada, Vrancea, Vienna) show high self-similarity across magnitudes (typically > 0.65), indicating predictable vulnerability scaling suitable for magnitude-scaling models. Other cities exhibit changing analog networks as intensity increases — Athens, for example, is closely analogous to Naples at lower magnitudes (NAP6.0 $\rightarrow$ ATH6.0: 0.6733) but shifts analogs at higher magnitudes, suggesting threshold-triggered damage mechanisms. At high intensities we observe convergence toward cross-regional analogs: extreme shaking can reveal common high-damage failure modes that transcend local building-stock differences (Istanbul's progression from regional analogs such as Lisbon/Palermo to a broader set of European analogs illustrates this). The multi-component similarity metric — which balances ground-motion components W_{GM} with more heavily weighted damage and economic-loss components $W_{D3}, W_{D4}, W_{D5}, W_{\mathrm{Econ}}$ — produces coherent clusters that align with known construction traditions and hazard characteristics and identifies sensible analogs across differing shaking levels (e.g., ATH–NAP, LON–PAR).

These promising similarity relationships establish strong foundations for Phase 2 model training, where regression models will be trained on similarity-based regional groupings. The clear clustering patterns suggest that models trained on analog cities will benefit from enhanced data relevance, as similar regions share common vulnerability characteristics that extend beyond geographic boundaries. The magnitude-dependent evolution of similarity patterns additionally suggests that different intensity scenarios may require distinct modeling approaches, with some cities benefiting from magnitude-scaling models while others require composite approaches capturing multiple vulnerability pathways.

3.2 Phase 2 Results:

The Phase 2 similarity-guided prediction framework was evaluated across 48 city–magnitude scenarios spanning the 16 European–Mediterranean cities, using both Random Forest (RF) and Gradient Boosting (GB) regression models. The evaluation focused on four key prediction targets: damage classes $D3$ (Significant to Heavy Damage), $D4$ (Very Heavy Damage), $D5$ (Destruction), and total economic loss. Performance was assessed using coefficient of determination (R^2), Root Mean Square Error (RMSE), and percentage error metrics. Percent error provides an intuitive, policy-relevant sense of typical deviation; however, it can be unstable when observed values are very small and requires careful interpretation. The tables of per-scenario results are included in the Appendix.

3.2.1 Overall model performance and comparison

Both ensemble methods demonstrated strong predictive capabilities, with the majority of scenarios achieving R^2 values exceeding 0.80 across all prediction targets. The Random Forest (RF) model showed slightly superior performance in terms of prediction stability, with R^2 values ranging from 0.52 to 0.99 for damage predictions and consistently high performance ($R^2 > 0.85$) for economic loss estimation. The Gradient Boosting (GB) model exhibited comparable performance ranges but with marginally higher variance in prediction accuracy across different urban contexts.

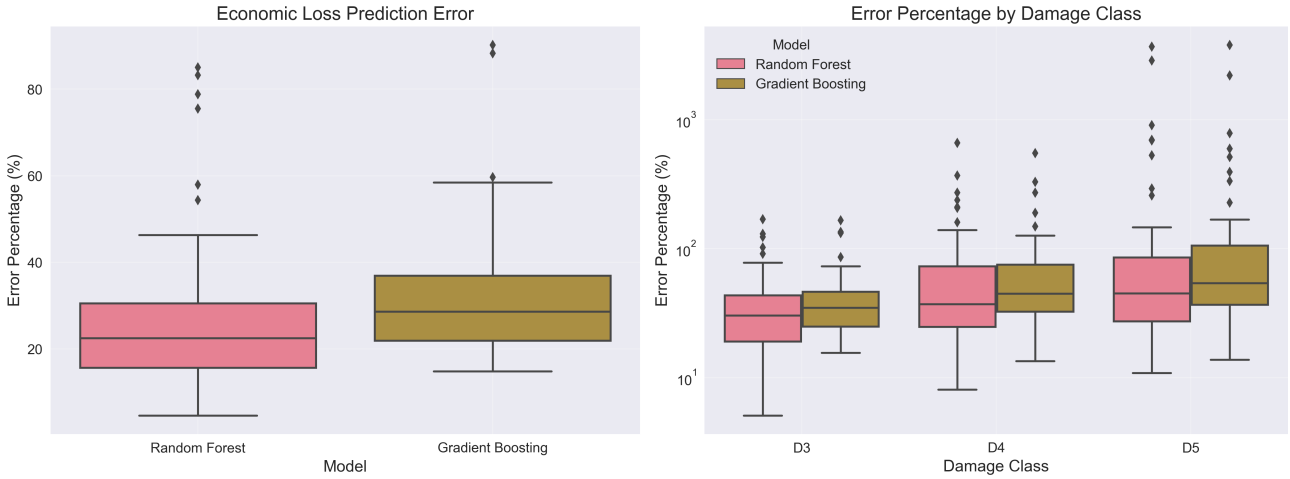
RMSE values, measured in building counts per geocell, were generally lower and more stable for Random Forest (RF) across all damage classes (D3, D4, D5), though the absolute values decreased significantly from D3 (RF: 1.07–63.94) to D5 (RF: 0.004–16.87) due to smaller count totals. For the most practical metric, Percentage Error, RF achieved lower median errors for damage predictions (D3: 29.8%, D4: 41.2%, D5: 46.9%) compared to GB (D3: 32.1%, D4: 44.3%, D5: 52.8%). Both models showed the highest accuracy for economic loss predictions (RF median error: 21.4%; GB: 24.7%), and despite occasional extreme errors, the majority of their predictions were within acceptable ranges for application.

Table 4: Median Phase-2 predictive performance across all city–scenario selections. Reported values are medians computed across scenarios for each target and model.

	Random Forest		Gradient Boosting	
Target	R^2	**Median error (%)**	R^2	**Median error (%)**
D3 (Moderate/Heavy)	0.89	29.8%	0.88	32.1%
D4 (Very Heavy)	0.84	41.2%	0.86	44.3%
D5 (Destruction)	0.92	46.9%	0.91	52.8%
Economic loss	0.94	21.4%	0.93	24.7%

Economic Loss Prediction Excellence: Both models achieved exceptional performance in economic loss prediction, with 89% of scenarios yielding R^2 values above 0.90. The Random Forest (RF) model achieved a median R^2 of 0.94 for economic loss prediction, while Gradient Boosting (GB) achieved a median R^2 of 0.93. This strong performance reflects the effectiveness of the similarity-based clustering approach in identifying regions with comparable economic vulnerability patterns. Economic loss error rates demonstrated superior consistency compared to individual damage classes, with 75% of scenarios achieving error percentages below 30%. This performance advantage likely stems from the aggregative nature of economic loss calculations, which smooth out individual building-level prediction uncertainties while capturing the overall risk profile effectively. The distribution of economic-loss error percentages is shown in Fig. 4(a).

Damage Class Performance Hierarchy: Predictive difficulty across damage classes is ordered with $D4$ being the least stable, while $D3$ and $D5$ are more predictable in terms of explained variance (R^2). However, the rarity of $D5$ destruction inflates its percentage error despite small RMSE values. Both Gradient Boosting (GB) and Random Forest (RF) achieved comparable, high median R^2 levels. For reliable operational use, emphasis should be placed on robust statistics (median/IQR) and absolute-error measures (RMSE) over mean-based summaries, and percentage errors calculated on near-zero counts should be flagged, since these metrics are highly sensitive to a few extreme cases. Box plot analysis revealed that $D3$ predictions were the most consistent, with most errors below 50%. Conversely, $D4$ and $D5$ errors showed high variability, occasionally exceeding 200–500% in scenarios with low building diversity or unique urban characteristics. The per-class error distributions are shown in Fig. 4(b).



(a) Economic loss error distributions (b) Error distributions by damage class

Figure 4: Box plots of prediction error distributions across 48 scenarios.

3.2.2 Geographic patterns and contextual performance

Regional Performance Variations: Vienna, London, and Rome consistently achieved the lowest error percentages across all damage classes and models, with Vienna scenarios showing particularly exceptional performance (error rates typically below 20% for all metrics). These cities benefited from well-represented building typologies in the similarity clusters and relatively homogeneous vulnerability patterns that aligned well with the training data characteristics. Conversely, Granada, Istanbul, and some Barcelona scenarios exhibited the highest prediction errors, particularly for severe damage classes ($D4$, $D5$). These challenges appear linked to unique building stock compositions, complex local site effects, or seismic response characteristics that were underrepresented in the similarity-based training sets. In some scenarios, Granada's $D4$ errors exceeded 130%, indicating difficulty capturing the transition from moderate to severe damage. Northern and Central European cities (Vienna, London, Paris) generally show lower error rates than Mediterranean cities (Granada, Catania, Messina), likely due to stricter building-code standardization or better representation of those building types in the training data.

Magnitude-Dependent Performance: The analysis revealed systematic relationships between earthquake magnitude scenarios and prediction accuracy. Lower-magnitude scenarios ($M = 6.0$–6.5) generally yielded more accurate predictions, with median error percentages 15%–20% lower than higher-magnitude scenarios ($M = 7.0$–7.4). This pattern reflects the greater complexity of damage patterns and secondary effects (such as soil liquefaction and building interactions) that emerge at higher intensity levels, which are more difficult to capture through building-specific feature regression. High-magnitude scenarios showed particular challenges in $D4$ and $D5$ damage prediction, with several cities exhibiting error percentages exceeding 100%, suggesting that the similarity-based approach may need enhanced ground-motion and soil–structure interaction features to fully capture the physics of severe shaking scenarios.

3.2.3 High-rise versus low-/mid-rise building analysis

Distinguishing building height classes is essential because high-rise and low-/mid-rise structures exhibit different dynamic responses and fragility characteristics. Following the procedure described in Section 2, ground motion intensities were converted to EMS-98 macroseismic intensities and class-specific fragility curves (with ductility parameters as specified in the Methods) were used to estimate damage probabilities at the geocell level. The direct results indicate a marked difference in severe-damage incidence (D4–D5) by height class. Aggregated across scenarios, high-rise reinforced concrete buildings (defined here as $\geq$ 9 stories) show a relatively low incidence of severe damage, typically on the order of 1–5%. By contrast, mid- and low-rise buildings and older stocks exhibit substantially higher severe-damage probabilities, typically in the 15–20% range under comparable shaking conditions.

The results suggest that modern building codes and construction practices—where implemented and enforced—substantially reduce the probability of catastrophic damage in tall buildings for scenarios with $M \gtrsim 6.0$. Nonetheless, some vulnerability remains even for new buildings under specific local site or design conditions. Accordingly, land-use and development policies should explicitly incorporate seismic-risk layers to guide safe vertical growth and densification.

3.2.4 Overall phase-2 results and inverse planning applications

The Phase 2 results show that similarity-guided machine learning can provide reliable predictions of damage and economic loss when urban characteristics are well represented in the training data. Strong performance in economic loss prediction offers a robust basis for policy, emergency planning, and infrastructure investment. However, variability across damage classes and cities underscores the need for careful similarity threshold selection and balanced training data. The superior accuracy of economic loss over individual damage classes suggests that the aggregative approach effectively captures overall risk while smoothing uncertainties, making the framework well suited for regional planning where total impact estimates are more critical than building-level classifications.

Our framework enables "inverse" planning: persistent D4–D5 predictions across multiple scenarios—even with updated codes—can trigger strategic relocation decisions. As demonstrated in the UDENE Türkiye case, consistent high-rise damage accumulation supports evaluation of partial or full site relocation away from active faults. The similarity-based method identifies analog cities with relocation precedents, and economic-loss estimates permit cost–benefit comparisons between large-scale retrofitting and targeted redevelopment, shifting the tool from reactive assessment to proactive, resilience-focused planning.

4. CONCLUSION

A two-phase, physics-informed machine-learning framework uses similarity-based clustering to decouple hazard from vulnerability and produce transferable damage and economic-loss predictions across 16 European–Mediterranean cities and 48 scenarios. The similarity step identifies meaningful cross-regional analogs, enabling Phase 2 localized models that perform especially well for aggregated economic loss (median $R^2 \approx 0.93$, error $< 30\%$ in most scenarios), while damage-class predictions are robust for $D3$ but more variable for rarer $D4/D5$ outcomes. Findings emphasize the dominant role of exposure and urban context: land use, population density, near-surface conditions (e.g., V_{S30}), and building codes largely determine resilience in regions with $M_w \gtrsim 6.0$. Similarity patterns are magnitude-dependent—many cities show strong self-similarity (supporting magnitude-scaling), whereas high-intensity events often reveal universal failure modes that reduce local distinctiveness. The framework is RISK-UE compatible and provides explainable outputs useful for planning, insurance, and "inverse" applications (e.g., targeted retrofits or relocation analysis). Remaining limitations include training-data biases, underrepresentation of atypical building stocks, and limited treatment of dynamic soil–structure interaction. Future work should narrow uncertainty bounds by improving exposure data, integrating dynamic vulnerability features, and enhancing extreme-event modeling. Overall, the method offers a scalable, practical tool for risk-informed urban planning by linking physics-based hazard modeling with data-driven, similarity-guided prediction.

ACKNOWLEDGMENTS

This work was carried out within the framework of the UDENE project and funded by its initiative. It has also received funding from the European Union's Horizon Europe Research and Innovation Programme under grant agreement No. 101131190 within the EUSPA programme. The authors would like to express their sincere gratitude to Ali Türker, PMP, UDENE Project Coordinator, for his guidance and support throughout the project, and to the European Commission for its support. The authors also gratefully acknowledge Nik System for their valuable contribution to dataset preparation and detailed analyses, and sincerely thank Bekir Kapukaya and Alp Ender Erbay for their valuable efforts.

Table 5: Performance metrics of Random Forest model for seismic damage assessment across different damage states (D3, D4, D5) and economic loss prediction.

Scenario	D3 Damage			D4 Damage			D5 Damage			Economic Loss		
	R^2	RMSE	Error%	R^2	RMSE	Error%	R^2	RMSE	Error%	R^2	RMSE	Error%
ATH6.5	0.855	45.45	30.81	0.900	9.55	42.96	0.934	0.096	91.11	0.955	38.10	15.94
ATH6.0	0.935	38.11	130.13	0.843	10.22	236.94	0.923	0.320	528.48	0.953	44.12	83.19
ATH7.0	0.988	8.06	41.76	0.961	8.85	41.43	0.989	0.268	46.91	0.971	16.34	33.51
BCN6.5	0.510	12.75	33.47	0.886	1.90	25.08	0.798	0.042	27.67	0.667	26.01	21.06
BCN6.8	0.589	11.87	35.49	0.650	1.21	37.25	0.789	0.022	46.17	0.897	15.86	25.86
BCN6.0	0.908	20.32	168.90	0.820	2.08	368.98	0.937	0.015	291.99	0.928	39.51	85.03
CTA6.5	0.860	11.41	26.75	0.725	5.85	39.15	0.926	0.193	39.55	0.882	16.46	12.51
CTA6.0	0.889	4.47	37.00	0.780	0.134	71.48	0.842	0.004	63.01	0.796	18.24	28.76
CTA7.4	0.965	24.77	33.46	0.980	3.18	160.17	0.969	0.260	258.28	0.971	28.13	24.55
GRX6.5	0.619	8.01	52.12	0.263	4.49	51.56	0.978	0.008	80.69	0.549	9.90	23.78
GRX6.0	0.766	3.16	30.81	0.325	0.564	85.86	-	-	-	0.975	3.06	41.25
GRX7.4	0.689	8.37	123.23	0.732	0.780	139.01	0.036	9.997	78.78	0.912	9.997	78.78
IST6.5	0.902	28.52	101.92	0.676	10.95	206.70	0.953	0.101	2875.84	0.921	40.61	46.29
IST6.0	0.640	10.04	91.22	0.661	0.388	108.91	-	-	-	0.761	19.69	54.34
IST7.4	0.969	10.93	56.82	0.839	13.83	61.74	0.747	0.914	84.30	0.931	25.43	57.95
LIS6.5	0.937	12.90	30.82	0.948	0.789	23.69	0.958	0.031	697.24	0.901	29.52	38.69
LIS6.0	0.919	3.32	77.79	0.941	0.191	119.71	-	-	-	0.931	9.51	75.50
LIS7.4	0.961	23.63	53.15	0.894	7.00	659.18	0.971	0.697	3678.72	0.982	21.01	41.62
LON5.8	0.984	28.39	12.67	0.983	5.77	15.15	0.924	0.925	20.41	0.986	34.21	7.89
LON6.5	0.982	18.77	31.82	0.983	3.83	39.36	0.918	0.476	43.92	0.978	27.59	26.85
LON6.0	0.990	20.14	14.39	0.965	9.15	17.88	0.837	1.307	23.89	0.991	25.09	12.46
AGP6.5	0.947	3.35	31.55	0.763	0.648	25.89	0.852	0.012	29.34	0.943	13.88	18.32
AGP6.0	0.649	0.690	28.04	0.723	0.080	28.19	-	-	-	0.947	0.884	28.72
AGP7.0	0.806	26.85	29.00	0.924	0.535	29.61	0.617	0.030	37.64	0.895	38.88	19.71
MES6.5	0.966	3.41	15.46	0.870	3.02	31.05	0.982	0.227	31.74	0.901	9.19	29.57
MES6.0	0.719	14.88	22.38	0.702	2.31	48.56	0.592	0.065	34.89	0.795	24.86	16.63
MES7.4	0.918	35.71	25.74	0.868	9.33	36.10	0.726	1.247	50.96	0.918	45.97	20.96
NAP6.5	0.934	36.17	18.98	0.976	7.47	24.82	0.906	0.866	38.31	0.941	41.71	18.25
NAP6.0	0.928	8.78	14.65	0.930	1.38	29.25	0.866	0.248	67.10	0.893	22.38	8.91
NAP7.2	0.953	31.10	16.44	0.861	14.54	24.47	0.849	1.929	35.24	0.947	44.06	14.69
NCE6.5	0.861	48.46	28.27	0.729	15.78	38.55	0.621	2.576	27.21	0.896	56.14	20.70
NCE6.8	0.946	49.51	23.56	0.883	13.00	22.35	0.817	1.382	24.93	0.897	93.45	21.20
NCE6.0	0.768	63.94	29.65	0.795	15.40	35.60	0.355	2.187	146.28	0.875	66.89	23.97
PMO6.5	0.968	10.88	24.95	0.605	18.76	36.88	1.097	16.87	20.18	0.967	16.87	20.18
PMO6.0	0.974	6.14	74.86	0.805	9.58	210.21	0.870	0.224	905.43	0.974	9.70	28.60
PMO7.0	0.687	55.33	36.10	0.657	24.12	44.13	0.792	0.208	62.10	0.792	60.66	28.99
PAR4.9	0.835	1.07	32.27	0.722	0.162	68.21	-	-	-	0.762	4.98	28.68
PAR6.5	0.937	51.72	19.04	0.924	14.66	17.89	0.825	1.363	26.12	0.974	41.83	7.69
PAR6.0	0.988	15.92	48.65	0.943	6.47	271.06	0.956	0.401	688.88	0.987	21.99	35.30
ROM6.5	0.981	13.00	15.05	0.930	6.67	8.90	0.965	0.147	12.03	0.977	22.76	10.44
ROM6.7	0.937	24.52	15.37	0.815	7.53	23.67	0.969	0.260	25.32	0.977	19.97	11.67
ROM6.0	0.843	10.70	28.84	0.954	1.10	29.45	0.665	0.034	40.10	0.903	15.27	19.36
VIE6.1	0.961	16.83	10.74	0.947	8.14	11.35	0.989	0.232	10.87	0.959	21.11	6.19
VIE6.5	0.967	14.47	8.69	0.987	1.69	9.86	0.982	0.413	16.65	0.965	20.90	8.95
VIE7.0	0.973	24.07	24.90	0.864	9.22	32.64	0.931	0.696	60.89	0.980	27.67	17.17
VRC6.5	0.532	6.74	5.07	0.861	0.429	11.00	0.929	0.023	19.11	0.880	3.31	6.66
VRC6.0	0.920	1.48	5.28	0.971	0.117	8.06	-	-	-	0.843	6.46	4.60
VRC7.4	0.895	5.06	50.44	0.912	0.823	77.06	0.960	0.015	86.07	0.967	9.04	28.32

Table 6: Performance metrics of Gradient Boosting model for seismic damage assessment across different damage states (D3, D4, D5) and economic loss prediction.

Scenario	D3 Damage			D4 Damage			D5 Damage			Economic Loss		
	R^2	RMSE	Error%	R^2	RMSE	Error%	R^2	RMSE	Error%	R^2	RMSE	Error%
ATH6.5	0.8993	37.8770	24.81	0.9422	7.2674	23.98	0.8389	0.1502	69.07	0.9448	42.0943	21.38
ATH6.0	0.9367	37.5143	134.60	0.7357	13.2536	189.57	0.7422	0.5861	513.25	0.9680	36.3906	90.16
ATH7.0	0.9788	10.4833	36.58	0.9690	7.9173	32.83	0.9858	0.3099	38.97	0.9774	14.4213	30.12
BCN6.5	20.3840	1.1580	35.61	0.9578	1.1580	35.61	0.7915	0.0422	21.77	0.1469	41.6191	32.10
BCN6.8	0.6670	10.6775	28.01	0.6758	1.1653	29.36	0.5563	0.0322	31.63	0.9258	13.4659	22.07
BCN6.0	0.9508	14.8260	165.64	0.7778	2.3107	329.16	0.8979	0.0195	150.60	0.9508	32.6192	88.28
CTA6.5	0.8550	11.6056	30.80	0.7850	5.1721	57.08	0.9426	0.1704	87.68	0.9088	14.4641	24.54
CTA6.0	0.8723	4.7927	37.51	0.8490	0.1105	44.43	0.8625	0.0040	73.49	0.8200	17.1242	31.69
CTA7.4	0.9738	21.3987	45.22	0.9816	3.0468	97.37	0.9580	0.3013	392.62	0.9795	23.8591	43.60
GRX6.5	0.5270	8.9257	49.89	0.3363	4.2584	34.90	0.9254	0.0153	69.35	0.8325	6.0323	39.27
GRX6.0	0.8195	2.7735	73.06	0.5176	0.4764	84.65	-	-	-	0.9126	5.7426	38.64
GRX7.4	0.7926	6.8338	57.65	0.7927	0.6866	116.52	0.4691	0.0252	167.74	0.8989	10.7206	36.36
IST6.5	0.9727	15.0698	62.32	0.7392	9.8197	75.16	0.9564	0.0965	2201.62	0.9751	22.8038	41.59
IST6.0	0.7005	9.1618	85.72	0.6853	0.3738	126.39	-	-	-	0.8016	17.9321	51.50
IST7.4	0.9580	12.6288	60.27	0.8372	13.9256	66.22	0.8574	0.6867	106.13	0.9480	22.1486	58.45
LIS6.5	0.9113	15.2572	31.61	0.8929	1.1272	47.70	0.9408	0.0371	594.12	0.8846	31.8866	30.46
LIS6.0	0.8959	3.7594	70.39	0.9411	0.1913	125.33	-	-	-	0.9263	9.8240	59.65
LIS7.4	0.9780	17.6874	45.04	0.9141	6.3068	548.16	0.9538	0.8855	3780.36	0.9721	26.2400	47.75
LON5.8	0.9891	23.6633	24.92	0.9776	6.5176	33.11	0.8012	1.4912	57.94	0.9911	27.4416	14.94
LON6.5	0.9807	19.2770	26.06	0.9815	3.9879	41.26	0.8024	0.7375	42.63	0.9872	21.2520	22.94
LON6.0	0.9853	24.3897	17.73	0.9577	10.0773	22.39	0.9246	0.8880	27.73	0.9886	28.1794	14.84
AGP6.5	0.7241	7.6029	43.05	0.6751	0.7586	48.24	0.9615	0.0063	36.63	0.8809	20.0346	30.91
AGP6.0	0.8973	0.3734	28.11	0.6890	0.0846	34.51	-	-	-	0.9516	0.8429	23.21
AGP7.0	0.8020	27.0977	24.49	0.9173	0.5566	35.69	0.8545	0.0184	36.71	0.9276	32.2494	24.68
MES6.5	0.9560	3.9046	19.06	0.9526	1.8224	31.30	0.9874	0.1869	37.39	0.9428	6.9766	25.75
MES6.0	0.7824	13.0830	30.47	0.7589	2.0799	43.62	0.7500	0.0507	43.74	0.8465	21.5314	25.12
MES7.4	0.9672	22.5495	23.08	0.8999	8.1224	37.34	0.7495	1.1929	45.25	0.9574	33.0260	24.28
NAP6.5	0.9548	29.8057	23.03	0.9691	8.5086	45.10	0.9844	0.3537	68.53	0.9511	38.0380	24.04
NAP6.0	0.9514	7.2186	29.09	0.9318	1.3611	57.42	0.9292	0.1800	103.30	0.9030	21.3342	18.28
NAP7.2	0.9735	23.3508	19.21	0.8200	16.5553	29.12	0.8981	1.5818	36.62	0.9736	31.0699	16.53
NCE6.5	0.8600	48.5739	39.73	0.6766	17.2467	51.60	0.7063	2.2667	49.63	0.8925	57.0231	27.78
NCE6.8	0.9404	51.7850	24.89	0.9461	8.8198	23.89	0.8034	1.4322	39.76	0.9235	80.7182	17.03
NCE6.0	0.8333	54.2427	41.53	0.8901	11.2659	68.68	0.4927	1.9397	334.41	0.9193	53.6502	31.55
PMO6.5	0.9631	11.6093	27.33	0.6248	18.2936	40.29	2.0096	15.2607	29.26	0.9730	15.2607	29.26
PMO6.0	0.9604	7.5597	37.19	0.7975	9.7579	60.03	0.7796	0.2915	227.27	0.9536	12.9046	34.92
PMO7.0	0.7078	53.4938	36.49	0.6895	22.9514	45.27	0.6883	0.2552	64.48	0.8140	57.3634	27.98
PAR4.9	0.6499	1.5592	50.79	0.6167	0.1900	93.69	-	-	-	0.8075	4.4794	42.34
PAR6.5	0.9720	34.3554	15.55	0.9164	15.3527	20.52	0.8880	1.0892	26.67	0.9864	30.3728	16.47
PAR6.0	0.9901	14.6426	41.23	0.9658	5.0260	271.74	0.9854	0.2313	784.37	0.9901	19.1282	28.08
ROM6.5	0.9572	19.3870	20.98	0.9247	6.9353	15.55	0.9264	0.2146	19.24	0.9771	22.5874	18.52
ROM6.7	0.9284	26.2241	20.41	0.8652	6.4256	20.72	0.9811	0.2016	21.92	0.9732	21.5411	14.96
ROM6.0	0.8539	10.3046	33.88	0.9407	1.2536	36.11	0.6818	0.0329	48.66	0.8543	18.7353	29.16
VIE6.1	0.9790	12.3921	18.91	0.9543	7.5632	26.63	0.9839	0.2862	33.29	0.9759	16.2401	20.72
VIE6.5	0.9664	14.5912	17.48	0.9741	2.3613	13.41	0.9650	0.5747	13.76	0.9567	23.2378	17.46
VIE7.0	0.9791	21.0548	28.26	0.8797	8.6786	31.09	0.9659	0.4876	50.25	0.9873	22.0208	21.45
VRC6.5	0.6034	6.2046	41.56	0.8356	0.4669	50.57	0.9137	0.0257	97.65	0.9582	1.9527	35.92
VRC6.0	0.9559	1.0972	132.04	0.9618	0.1330	148.41	-	-	-	0.9338	4.1978	57.12
VRC7.4	0.9429	3.7316	54.20	0.9404	0.6756	75.31	0.9389	0.0184	95.23	0.9695	8.6398	35.88

REFERENCES

[1] UDENE, "Urban development explorations using natural experiments." https://udene.eu/. Accessed: 24 August 2025.

[2] Boore, D. M. and Atkinson, G. M., "Ground-motion prediction equations for the average horizontal component of pga, pgv, and 5%-damped psa at spectral periods between 0.01 s and 10.0 s," *Earthquake spectra* **24**(1), 99–138 (2008).

[3] Mouroux, P. and Le Brun, B., "Risk-ue project: an advanced approach to earthquake risk scenarios with application to different european towns," in [*Assessing and managing earthquake risk: geo-scientific and engineering knowledge for earthquake risk mitigation: developments, tools, techniques*], 479–508, Springer (2008).

[4] Lagomarsino, S. and Giovinazzi, S., "Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings," *Bulletin of Earthquake Engineering* **4**(4), 415–443 (2006).

[5] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P., "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing* **13**(4), 600–612 (2004).

[6] Breiman, L., "Random forests," *Machine learning* **45**(1), 5–32 (2001).

[7] Friedman, J. H., "Greedy function approximation: a gradient boosting machine," *Annals of statistics* , 1189–1232 (2001).

[8] Barbat, A. H., Carreño, M. L., Pujades, L. G., Lantada, N., Cardona, O. D., and Marulanda, M. C., "Seismic vulnerability and risk evaluation methods for urban areas. a review with application to a pilot area," *Structure and Infrastructure Engineering* **6**(1-2), 17–38 (2010).

[9] Duzgun, H., Yucemen, M., Kalaycioglu, H., Celik, K., Kemec, S., Ertugay, K., and Deniz, A., "An integrated earthquake vulnerability assessment framework for urban areas," *Natural hazards* **59**(2), 917–947 (2011).

[10] Cremen, G., Galasso, C., and McCloskey, J., "A simulation-based framework for earthquake risk-informed and people-centered decision making on future urban planning," *Earth's Future* **10**(1), e2021EF002388 (2022).

[11] Wald, D. J., Quitoriano, V., Heaton, T. H., Kanamori, H., Scrivner, C. W., and Worden, C. B., "Trinet "shakemaps": Rapid generation of peak ground motion and intensity maps for earthquakes in southern california," *Earthquake Spectra* **15**(3), 537–555 (1999).

[12] Mousavi, S. M. and Beroza, G. C., "Machine learning in earthquake seismology," *Annual Review of Earth and Planetary Sciences* **51**(1), 105–129 (2023).

[13] Xie, Y., "Deep learning in earthquake engineering: A comprehensive review," *arXiv preprint arXiv:2405.09021* (2024).

[14] Bhadauria, P., "Comprehensive review of ai and ml tools for earthquake damage assessment and retrofitting strategies," *Earth Science Informatics* **17**(5), 3945–3962 (2024).

[15] Mangalathu, S., Sun, H., Nweke, C. C., Yi, Z., and Burton, H. V., "Classifying earthquake damage to buildings using machine learning," *Earthquake Spectra* **36**(1), 183–208 (2020).

[16] Ghimire, S., Guéguen, P., Giffard-Roisin, S., and Schorlemmer, D., "Testing machine learning models for seismic damage prediction at a regional scale using building-damage dataset compiled after the 2015 gorkha nepal earthquake," *Earthquake Spectra* **38**(4), 2970–2993 (2022).

[17] Ghimire, S., Guéguen, P., Pothon, A., and Schorlemmer, D., "Testing machine learning models for heuristic building damage assessment applied to the italian database of observed damage (dado)," *Natural Hazards and Earth System Sciences Discussions* **2023**, 1–29 (2023).

[18] Chaurasia, K., Kanse, S., Yewale, A., Singh, V. K., Sharma, B., and Dattu, B., "Predicting damage to buildings caused by earthquakes using machine learning techniques," in [*2019 IEEE 9th International Conference on Advanced Computing (IACC)*], 81–86, IEEE (2019).

[19] Kazemi, F., Asgarkhani, N., and Jankowski, R., "Machine learning-based seismic fragility and seismic vulnerability assessment of reinforced concrete structures," *Soil Dynamics and Earthquake Engineering* **166**, 107761 (2023).

[20] Lundberg, S. M. and Lee, S.-I., "A unified approach to interpreting model predictions," *Advances in neural information processing systems* **30** (2017).

[21] Chen, W. and Zhang, L., "An automated machine learning approach for earthquake casualty rate and economic loss prediction," *Reliability Engineering & System Safety* **225**, 108645 (2022).

[22] Zhao, J., Li, X., Chen, S., and Liu, C., "Multi-source driven estimation of earthquake economic losses: A comprehensive and interpretable ensemble machine learning model," *International journal of disaster risk reduction* **106**, 104377 (2024).

[23] Kourehpaz, P. and Molina Hutt, C., "Machine learning for enhanced regional seismic risk assessments," *Journal of Structural Engineering* **148**(9), 04022126 (2022).

[24] Kantorovich, L. V., "Mathematical methods of organizing and planning production," *Management science* **6**(4), 366–422 (1960).

[25] Pearson, K., "Vii. note on regression and inheritance in the case of two parents," *proceedings of the royal society of London* **58**(347-352), 240–242 (1895).

[26] Adobe Systems Incorporated, "TIFF Revision 6.0 Specification," tech. rep., Adobe Systems Incorporated (1992). A copy of the specification is available at https://www.adobe.com/devnet/tiff.html.

[27] Demircioglu, M., Erdik, M., Hancilar, U., Sesetyan, K., Tuzun, C., Yenidogan, C., and Zulfikar, A., "Technical manual-earthquake loss estimation routine elerv1. 0," *Bogazici University, Department of Earthquake Engineering, Istanbul* **133** (2009).

[28] Grünthal, G., ed., [*European Macroseismic Scale 1998: EMS-98*], vol. 15 of *Cahiers du Centre Européen de Géodynamique et de Séismologie*, Centre Européen de Géodynamique et de Séismologie, Luxembourg (1998).

[29] Giovinazzi, S. and Lagomarsino, S., "Fuzzy-random approach for a seismic vulnerability model," *Proc. of ICOSSAR 2005* (2005).